



NQF
NATIONAL QUALITY FORUM



**COMPOSITE
FRAMEWORK
AND
MEASURES**



Composite Measure Evaluation Framework
and National Voluntary Consensus Standards
for Mortality and Safety—Composite Measures

A CONSENSUS REPORT

COMPOSITE EVALUATION FRAMEWORK STEERING COMMITTEE

Harlan Krumholz, MD (Co-Chair)

Yale University School of Medicine
New Haven, CT

Jinnet Briggs Fowles, PhD (Co-Chair)

Park Nicollet Institute
St. Louis Park, MN
(resigned from Steering Committee in 2007)

Gail Amundson, MD

Quality Quest for Health of Illinois
Peoria, IL

Elizabeth DeLong, PhD

Duke University Clinical Research Institute
Durham, NC

Margaret Foley, PhD

Temple University
Philadelphia, PA

Ted von Glahn, MS

Pacific Business Group on Health
San Francisco, CA

Arthur Pelberg, MD, MPA

Schaller Anderson
Phoenix, AZ

David Shahian, MD

Massachusetts General Hospital
Boston, MA

Steven Wright, PhD

Veterans Health Administration
Providence, RI

Liaison Member

Through August 2008

Marybeth Farquhar, RN, MSN, PhD

Agency for Healthcare Research and Quality
Rockville, MD

PROJECT STAFF

Helen Burstin, MD, MPH

Senior Vice President
Performance Measures

Lisa Hines, BSN, MS

Consultant

Karen Beckman Pace, PhD, RN

Senior Director

Melinda L. Murphy, RN, MS, NE-BC

Consultant

Composite Measure Evaluation Framework and National Voluntary Consensus Standards for Mortality and Safety—Composite Measures: A Consensus Report

Foreword

AS THE NATION MOVES FORWARD to improve its healthcare systems, performance measurement is fundamental to understanding healthcare’s current state and is critical to developing solutions for improvement. The criteria used to evaluate performance measures by the National Quality Forum (NQF) are an important part of the process used to ensure that NQF-endorsed® measures remain the gold standard for measuring healthcare quality.

NQF has long used a standard set of criteria for evaluating performance measures and has continued to refine the criteria as the need to do so is discovered through their use. The evolution and use of composite measures has presented just such a need. Combining measures of performance to convey a broader picture than can be done with single measures holds promise for improving understanding and stimulating improvement, but only if such measures are deemed to meet nationally accepted standards. The composite evaluation framework includes a set of criteria that were adapted from and build on the proven NQF performance measure evaluation criteria. Over two periods of review, healthcare stakeholders provided comments that helped to refine the framework and criteria. To test them, the criteria then were applied to a group of composite measures.

NQF thanks the members of the Composite Evaluation Framework Steering Committee and NQF Members for their commitment to ensuring that measures used to evaluate the performance of the nation’s healthcare organizations meet the criteria of importance, scientific acceptability, usability, and feasibility, using relevant constructs.



Janet M. Corrigan, PhD, MBA
President and Chief Executive Officer

The mission of the National Quality Forum is to improve the quality of American healthcare by setting national priorities and goals for performance improvement, endorsing national consensus standards for measuring and publicly reporting on performance, and promoting the attainment of national goals through education and outreach programs.

Recommended Citation: National Quality Forum (NQF). *Composite Measure Evaluation Framework and National Voluntary Consensus Standards for Mortality and Safety—Composite Measures: A Consensus Report*. Washington, DC: NQF; 2009.

This work was conducted under a grant from the Robert Wood Johnson Foundation (www.rwjf.org) and with funding by the Agency for Healthcare Research and Quality (www.ahrq.gov).

© 2009. *National Quality Forum*
All rights reserved

ISBN 978-1-933875-38-5

No part of this report may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means electronic, mechanical, photocopying, recording, or otherwise, without prior permission of the National Quality Forum. Requests for permission to reprint or make copies should be directed to:

Permissions
National Quality Forum
601 13th Street NW
Suite 500 North
Washington, DC 20005
Fax 202-783-3434
www.qualityforum.org

Composite Measure Evaluation Framework and National Voluntary Consensus Standards for Mortality and Safety—Composite Measures: A Consensus Report

Table of Contents

Executive Summary	v
Background	1
Purpose	2
Composite Measures	2
Definition of Composite Measure	2
Figure 1. Illustration of the Components of a Composite Measure	3
Types of Composite Measures	4
Evaluation Framework	5
Principles for Composite Measure Evaluation	5
Importance to Measure and Report	6
Scientific Acceptability of Measure Properties	6
Usability	7
Feasibility	8
Table 1. Individual and Composite Measure Evaluation Criteria	8
National Voluntary Consensus Standards for Mortality and Safety: Composite Measures ..	18
Table 2. National Voluntary Consensus Standards for Mortality and Safety:	
Composite Measures	19
Endorsed Composite Measures	20
Composite Measure Not Endorsed	21
Notes	22
Appendix A— Specifications of the National Voluntary Consensus Standards for Mortality and Safety—Composite Measures	A-1
Appendix B— Glossary	B-1

Composite Measure Evaluation Framework and National Voluntary Consensus Standards for Mortality and Safety—Composite Measures: A Consensus Report

Executive Summary

HEALTHCARE IS A COMPLEX AND multidimensional activity, and measurement of its quality should reflect that fact. Individual measures can provide important information, but there also is value in summarizing performance by combining the information from multiple measures. If done well, such a summary can convey quality from many different perspectives. The Institute of Medicine's 2006 report on performance measurement, *Performance Measurement: Accelerating Improvement*, noted that composite measures can enhance measurement to extend beyond tracking performance on separate measures and provide a potentially deeper view of the reliability of the care system.

A composite measure is a combination of two or more individual measures in a single measure that results in a single score. The National Quality Forum's (NQF's) measure evaluation criteria (importance to measure and report, scientific acceptability of measure properties, usability, and feasibility) apply to all types of measures being considered for NQF endorsement, although they were developed primarily for individual measures. With the increasing interest in composite measures, the Composite Evaluation Framework Steering Committee was appointed to address the additional considerations that are specifically relevant to evaluating such measures.

This document provides background, rationale, and evaluation criteria for composite measures; it builds on the NQF measure evaluation criteria. The intent is to provide guidance for NQF committees, Members, and measure developers and to make transparent how composite measures will be evaluated in the NQF process. Additionally, it endorses, for public reporting, three composite measures that were assessed using the evaluation criteria. Of note, reexamination of the component measures included in the composites was not carried out by the Composite Evaluation Framework Steering Committee because the component measures had been evaluated by other NQF committees and were either NQF endorsed[®] or recommended as components of the composite by the appropriate Technical Advisory Panel under the NQF Hospital Care 2007 project.

National Voluntary Consensus Standards for Mortality and Safety: Composite Measures

- Mortality for selected conditions
- Pediatric patient safety for selected indicators
- Patient safety for selected indicators

Composite Measure Evaluation Framework and National Voluntary Consensus Standards for Mortality and Safety—Composite Measures: A Consensus Report

Background

QUALITY MEASUREMENT is seen as fundamental to improving healthcare and ultimately health. As of May 2008, there were 376 NQF-endorsed[®] quality measures, including only a few composite measures. Many of the endorsed quality measures focus on a single care process, or even one step in the care process. However, healthcare is a complex and multidimensional activity, and measurement of its quality should reflect that fact. Ideally, assessment of healthcare quality should address the structures, processes, and outcomes of care, as well as the six Institute of Medicine (IOM) aims or goals for care provided in the healthcare system (safe, timely, effective, efficient, equitable, and patient centered). Individual measures can provide important information, but there also is value in summarizing performance by combining the information from multiple measures. If done well, such a summary can convey quality from many different perspectives. The 2006 IOM report on performance measurement, *Performance Measurement: Accelerating Improvement*, noted that composite measures can enhance measurement to extend beyond tracking performance on separate measures and can provide a potentially deeper view of the reliability of the care system.¹ In addition, such an approach can support efforts to rate providers by the quality of their care.

NQF used standard evaluation criteria to assess measures that are under consideration as potential consensus standards for their suitability as national voluntary consensus standards. The standard measure evaluation criteria (importance to measure and report, scientific acceptability of measure properties, usability, and feasibility) were recently reviewed and updated; however, composite measures were not specifically addressed. Thus, there is a need to specify standards for the assessment of composite measures.

As with all quality measures, careful design and evaluation of composite measures is imperative. There are unique issues introduced by composite methodology that require additional scrutiny, including the validity of the component measures; the methods used for scoring/aggregating and weighting the components; and issues involved in the interpretation of the composite score. Both the composite and its component measures need to be evaluated to determine the suitability of the composite for endorsement as a voluntary consensus standard.

Purpose

The purpose of this project was to identify a framework for evaluating composite measures. This report builds on and adds to the NQF measure evaluation criteria to specifically address the evaluation of composite measures. The intent is to provide guidance for NQF committees, Members, and measure developers and make transparent how composite measures will be evaluated within the NQF process. As part of this work, four composite measures previously submitted and held until a framework was developed were evaluated, in part as a test of the framework.

Composite Measures

Definition of Composite Measure

The definition of a composite measure is as follows:

A composite measure is a combination of two or more individual measures in a single measure that results in a single score.

Examples of composite measures from other fields include those related to intelligence and personality tests and various economic indexes. Some instruments or scales, such as the CAHPS® patient experience survey, with subscale domains and scores, also are considered a type of composite measure. However, in the case of scales and subscales, the subscales may not necessarily stand alone as individual measures.

An example of an NQF-endorsed composite quality measure is CAD: Optimally Managed Modifiable Risk (HealthPartners, NQF# 0076).

Description: Percentage of members who have optimally managed modifiable risk factors (LDL, tobacco non-use, blood pressure control, aspirin usage).

Numerator: All members from the denominator who reach treatment targets for all numerator components:

- Low-Density Lipoprotein (LDL) Screening—Coronary artery disease (CAD) population who had an LDL during the measurement year or the year prior to the measurement year with a level less than 100 for the most recent screening.
- Tobacco Non-User—CAD population with documented non-smoking status.
- Blood Pressure Control—CAD population whose blood pressure is in control less than 140/90 during the measurement year.
- Aspirin Usage—CAD population eligible for aspirin use who were on aspirin therapy.

Denominator: Members between 18 and 75 years of age as of December 31st of the reporting year, who were continually enrolled with not more than one-month break in coverage and have a diagnosis of coronary artery disease (CAD).

Some individual measures are used together as paired measures, and some individual measures include multiple steps in a single care process, but paired measures or measures with multiple steps in a single care process are not considered composite measures.

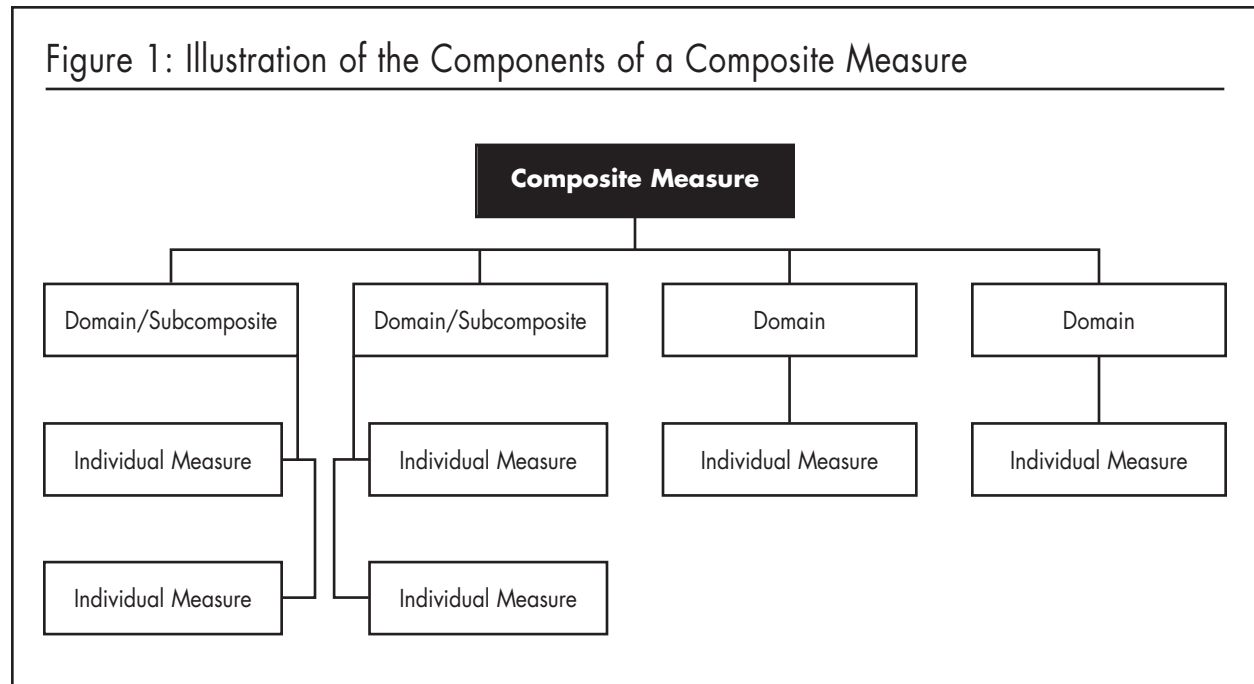
- Paired measures are individual measures that should be measured concurrently in the same population; however, the results are not combined into a single score (e.g., measuring mortality and readmission and displaying them together—but not calculating some joint score).

■ Individual steps in a multistep care process (i.e., assess, plan, treat), if addressed individually, generally should be included in one individual measure (e.g., assess immunization status and vaccinate eligible patients in one influenza immunization measure). The individual measure should focus on the step with the greatest effect on the desired outcome.

A variety of terms have been used to refer to composite measures and the components that comprise the composite measure. For this report, the terms *composite measure*,ⁱ *domain*

or *subcomposite*,ⁱⁱ and *individual measure*ⁱⁱⁱ are used to refer to the various levels and components.

Composites may be relatively simple, composed of only individual measures, or more complex, with various combinations of individual measures and composite measures. The illustration in Figure 1 depicts the components of a composite measure, including individual measures as well as domains or subcomposites that are composed of multiple measures.



ⁱ Other terms used for “composite measure” include *composite index*, *composite indicator*, *summary score*, *summary index*, and *scale*.

ⁱⁱ Other terms used for “domain or subcomposite” include *component composite*, *component index*, *factor*, *dimension*, and *subscale*.

ⁱⁱⁱ Other terms used for “individual measure” include *indicator*, *component measure*, *item*, and *variable*.

Many of the principles of composite measure development originated in both classical test theory and modern measurement theory and were initially applied to the social sciences, education, intelligence testing, and psychology. Those contemplating the development of composite measures should be familiar with the basic approaches used in these areas.

Constructing a composite measure entails the following major steps:^{2,9}

- Identify the purpose (e.g., comprehensive assessment of adult cardiac surgery quality of care) and delineate the quality construct to be measured (e.g., four domains of cardiac surgery quality include perioperative medical care, operative care, operative mortality, and postoperative morbidity).⁸
- Select the individual measures and/or subcomposite measures to be combined in the composite measure.
- Ensure that the weighting and scoring of the components supports the goal that is articulated for the measure. (Should the component scores be given equal weight or differential weights based on some prioritization? Should the component scores be standardized to achieve uniform scaling and/or directionality [i.e., if the component scores are on different scales, such as percentage of cardiac surgery patients on aspirin at discharge versus risk-adjusted operative mortality, or different direction, such as higher score on percentage of patients on aspirin reflects better quality versus higher risk-adjusted mortality reflects poorer quality]?).
- Combine the component scores, using a specified method, into one composite (e.g., sum, average, weighted average, patient-level all-or-none scoring).
- Finally, as with all measures, test the composite to determine if it is a reliable and valid indicator of quality healthcare.

Types of Composite Measures

There is no specific classification of composite measures that parallels those used for the three types of individual quality measures (structure, process, and outcome). Composite measures entail the combination of any number of various types of individual quality measures, although they ideally should reflect some common underlying construct such as quality of surgical care or quality of diabetes care. Composite measures often are described based on the method used to combine the component scores. Various methods may be used to combine the component scores into a composite (e.g., all-or-none scoring, sum, average, weighted average, opportunity scoring).

Many methods for combining the component measures use the provider-level results of the component measures (e.g., percentage of patients who are on aspirin at discharge and percentage of patients who received discharge beta blockade) to calculate the composite score. However, with all-or-none scoring, the composite result is determined at the patient level (e.g., percentage of patients who received all four medications: preoperative beta blockade, discharge aspirin, discharge beta blockade, and discharge antilipid therapy).^{10,11} Opportunity scoring is used with process measures and is determined from the sum of all numerators (i.e., number who achieved the desired process) divided by the sum of all denominators (i.e., number of eligible patients or opportunities, which could vary by measure).⁶

The Composite Evaluation Framework Steering Committee determined that many types of composite measures could be useful. Therefore, a broad definition of a combination of two or more individual measures was used to determine appropriate evaluation criteria.^{iv} In the evaluation criteria, no particular scoring/aggregation methodology is preferred. The methods used to develop and test the composite must be justified.

Evaluation Framework

The evaluation framework for composite measures follows NQF's standard evaluation criteria: importance to measure and report, scientific acceptability of measure properties, usability, and feasibility. These criteria are used for the individual measure components of the composite and also are relevant to the composite measure. The measure evaluation criteria and how they are applied to composite measures appear in Table 1. Some special considerations for composite measures that need to be addressed during evaluation include the need to a) standardize scores of the components if they have different scales or directionality; b) determine whether the components should be weighted differently and for what reason; and c) identify whether the scoring method is appropriate.

^{iv} The IOM report *Performance Measurement: Accelerating Improvement* identified composites as “denoting, at minimum, the combining of dichotomous indicators for several specific measures into a single number” and recommended an approach “to determine whether all critical aspects of care for a given condition have been achieved for an individual patient.” The purpose of this document is not to advocate for any particular approach to composite measures; rather it is to provide a framework for evaluating all of the types of composite measures that may be submitted to NQF.

Principles for Composite Measure Evaluation

Before identifying the specific evaluation criteria for composite measures, the Steering Committee articulated some general principles that underlay the evaluation of composite measures. These principles are addressed in the specific evaluation criteria identified in Table 1.

- The components of the composite (i.e., individual measures or component composite measures) must be either NQF-endorsed measures or determined to meet the individual measure evaluation criteria as the first step in evaluating the composite measure. A component measure might not be important enough in its own right as an individual measure, but it could be determined to be an important component of a composite. (This does not apply to subscales of scales/instruments that cannot be used independently of the total scale.)
- Even though all of the component measures must individually meet evaluation criteria, the composite measure as a whole also must meet evaluation criteria.
- Composites may be developed beginning with a conceptual construct of quality or with a set of measures one wishes to summarize into one score. The methods used to develop and test the components must be justified.

- Methods for combining the component scores influence the interpretation of the composite measure results and must be justified (e.g., all-or-none scoring indicates whether patients receive all/less than all of the items measured; averaging across component scores may obscure low or high scores of individual components).

Some principles that also apply to individual measures were highlighted for evaluating composite measures.

- Although composite measures result in a quantifiable score and use a variety of objective analytic methods, many subjective decisions influence the measure results, and the rationale for the chosen methods needs to be justified.
- The methods for constructing a composite should be explicitly stated and transparent so that the composite can be deconstructed.
- Combining multiple measures into a composite increases complexity, and using methodologically sound methods is of paramount importance. However, any additional methodological complexity that may be required should be fully transparent. Furthermore, even though the background calculations may be more complex, the final composite result should be simple and readily interpretable by all stakeholders.
- Ultimately, the justification for the composite measure is found in its effectiveness in accomplishing its intended purpose for the composite measure (i.e., to assess, and ultimately improve, the quality of healthcare).

Importance to Measure and Report

For the most part, if the component measures have been assessed and found to be sufficiently important to measure and report, then the composite will meet that criterion. In developing the framework, additional subcriteria were added for composite measures that relate to the purpose of creating a composite and the conceptual approach for selecting the components that make up the composite measure. In addition to being suitable for both public reporting and quality improvement, the purpose of creating a composite score (e.g., simplify the performance information presentation, identify whether all critical aspects of care were achieved) and the construct of quality should be described. Whether composite development begins with a conceptual construct of quality or with a set of measures one wishes to combine, the selection of the component measures should be conceptually coherent. The omission of important components that are indicated by the quality construct and purpose of measurement also could lead to validity problems and ultimately to difficulty in determining how to interpret the results of the composite score.

Scientific Acceptability of Measure Properties

Many of the subcriteria for individual measures also apply to composite measures, and each composite measure also must be tested to determine its scientific acceptability (e.g., whether it is a reliable and valid indicator of quality).^{7,12} Composite measures must be

precisely specified. In addition to the individual measure specifications, composite specifications include methods for standardizing scales from the various component scores, scoring rules (i.e., how the component scores are combined or aggregated), weighting rules (i.e., whether all component scores are given equal or differential weighting when combined into the composite), rules for the handling of missing data, and requirements for minimum case volumes for the component scores. The subcriteria for measure exclusions and risk adjustment apply only to the individual measures. Subcriteria are added to evaluate whether the components and scoring for a composite measure fit the conceptual model for the composite.

Several approaches might be used to combine measures. One traditional approach is the psychometric approach, developed in psychological and educational testing to create a measure of a complex construct that is not directly measurable, using multi-item scales.³ With the psychometric approach, the component items or measures are generally measuring the same underlying construct and should be correlated with one another (although not perfectly, or they would be redundant).¹³ Some composite measures may not reflect this classic psychometric construct, depending on the types of items or measures that are included in the composite.³ When the components are not correlated, the rationale and justification for their inclusion must be provided and appropriate analyses identified.

Decisions about weighting the various components of the composite measure can be based on expert opinion or empirical analyses (including factor analysis, principal components

analysis, and item response theory) and also should be consistent with the conceptual construct. If differential weighting is utilized, the weighting approach needs to be described and the rationale explained.

The approach to scoring and aggregation also needs to be described and the rationale explained. It should be consistent with the purpose and conceptual construct for the composite measure. For example, if the purpose is to determine if, and increase the likelihood that, all critical elements of care were provided, then an all-or-none scoring approach may be specified. Some scoring approaches, such as averages across multiple component scores, may compensate or obscure quality problems in one or more components—that is, a good score in one area may compensate for a poor score in another area.

Usability

As with individual measures, the main issue regarding the usability of a composite measure is whether the intended audiences find the information produced by the composite measure meaningful, understandable, and useful for both public reporting and quality improvement. It is critical that a composite measure, when reported, is readily decomposable into its constituent domains and individual measures. This will focus and facilitate quality improvement activities by providers and increase transparency and understanding of the measure results by all potential audiences. Additionally, it should be demonstrated that the purpose of creating a composite measure was achieved.

Feasibility

The first hurdle of feasibility is determining whether the component measures are considered feasible. Composites are more complex than individual measures, and in the case of all-or-none scoring, they require modifications

to the data collection methods used for individual measures in order to link individual data elements for individual patients. Therefore, the data collection strategy for obtaining all required components that need to be combined in the composite measure should be demonstrated to be feasible.

Table 1: Individual and Composite Measure Evaluation Criteria

The criteria for individual measure evaluation were updated with input from NQF Members, the public, and NQF's Consensus Standards Approval Committee and were approved by the NQF Board of Directors in August 2008.

INDIVIDUAL MEASURE EVALUATION CRITERIA	COMPOSITE MEASURE EVALUATION CRITERIA
<p>Conditions for Consideration</p> <p>Four conditions must be met before proposed measures may be considered and evaluated for suitability as voluntary consensus standards:</p> <ul style="list-style-type: none"> A. The measure is in the public domain or an intellectual property agreement is signed. B. The measure owner/steward verifies there is an identified responsible entity and process to maintain and update the measure on a schedule that is commensurate with the rate of clinical innovation, but at least every 3 years. C. The intended use of the measure includes both public reporting and quality improvement. D. The requested measure submission information is complete. Generally, measures should be fully developed and tested so that all of the evaluation criteria have been addressed and the information needed to evaluate the measure is provided. Measures that have not been tested are only potentially eligible for a time-limited endorsement, and, in that case, measure owners must verify that testing will be completed within 24 months of endorsement. 	<p>Conditions for Consideration</p> <p>The conditions for consideration of individual measures (A, B, C, D) also must be met for a composite measure.</p>

more

Table 1: Individual and Composite Measure Evaluation Criteria

The criteria for individual measure evaluation were updated with input from NQF Members, the public, and NQF's Consensus Standards Approval Committee and were approved by the NQF Board of Directors in August 2008.

INDIVIDUAL MEASURE EVALUATION CRITERIA	COMPOSITE MEASURE EVALUATION CRITERIA
<p>Criteria for Evaluation</p> <p>If all four conditions for consideration are met, measures are evaluated for their suitability based on four sets of standardized criteria: importance to measure and report, scientific acceptability of measure properties, usability, and feasibility. Not all acceptable measures will be strong—or equally strong—among each set of criteria. The assessment of each criterion is a matter of degree; however, all measures must be judged to have met the first criterion, importance to measure and report, in order to be evaluated against the remaining criteria.</p>	<p>Criteria for Evaluation</p> <p>The individual measures included in the composite or subcomposite measures must be either:</p> <p>NQF endorsed;</p> <p>OR</p> <p>assessed to have met the individual measure evaluation criteria as the first step in evaluating the composite measure.</p> <p>(This does not apply to subscales of a scale/ instrument that cannot be used independently of the total scale.)</p> <p>Following are the criteria that apply specifically to composite measure evaluation.</p>
<p>1. Importance to measure and report: Extent to which the specific measure focus is important to making significant gains in healthcare quality (safety, timeliness, effectiveness, efficiency, equity, patient-centeredness) and improving health outcomes for a specific high-impact aspect of healthcare where there is variation in or overall poor performance. Measures must be judged to be important to measure and report in order to be evaluated against the remaining criteria.</p> <p>1a. The measure focus addresses: a specific national health Goal/Priority identified by the Partners of the NQF-convened National Priorities Partnership</p> <p>OR</p>	<p>1. Importance to measure and report</p> <p>If the component measures are determined to meet the importance criteria 1a, 1b, and 1c, then the composite would meet 1a, 1b, and 1c. A component measure might not be important enough in its own right as an individual measure, but it could be determined to be an important component of a composite.</p>

more

Table 1: Individual and Composite Measure Evaluation Criteria

The criteria for individual measure evaluation were updated with input from NQF Members, the public, and NQF's Consensus Standards Approval Committee and were approved by the NQF Board of Directors in August 2008.

INDIVIDUAL MEASURE EVALUATION CRITERIA	COMPOSITE MEASURE EVALUATION CRITERIA
<p>a demonstrated high-impact aspect of healthcare (e.g., affects large numbers, leading cause of morbidity/mortality, high resource use [current and/or future], severity of illness, and patient/societal consequences of poor quality).</p> <p>1b. Demonstration of quality problems and opportunity for improvement, i.e., data¹ demonstrating considerable variation, or overall poor performance, in the quality of care across providers and/or population groups (disparities in care).</p> <p>1c. The measure focus is:</p> <p>an outcome (e.g., morbidity, mortality, function, health-related quality of life) that is relevant to, or associated with, a national health goal/priority, the condition, population, and/or care being addressed²;</p> <p>OR</p> <p>if an intermediate outcome, process, structure, etc., there is evidence³ that supports the specific measure focus as follows:</p> <p>Intermediate outcome – evidence that the measured intermediate outcome (e.g., blood pressure, Hba1c) leads to improved health/avoidance of harm or cost/benefit.</p> <p>Process – evidence that the measured clinical or administrative process leads to improved health/avoidance of harm and if the measure focus is on one step in a multistep care process,⁴ it measures the step that has the greatest effect on improving the specified desired outcome(s).</p>	<p>New for composite. 1d. The purpose/objective of the composite measure and the construct for quality are clearly described.</p> <p>New for composite. 1e. The component items/measures (e.g., types, focus) that are included in the composite are consistent with and representative of the conceptual construct for quality represented by the composite measure. Whether the composite measure development begins with a conceptual construct or a set of measures, the measures included must be conceptually coherent and consistent with the purpose.</p> <p><i>If not important to measure and report, STOP.</i></p>

more

Table 1: Individual and Composite Measure Evaluation Criteria

The criteria for individual measure evaluation were updated with input from NQF Members, the public, and NQF's Consensus Standards Approval Committee and were approved by the NQF Board of Directors in August 2008.

INDIVIDUAL MEASURE EVALUATION CRITERIA	COMPOSITE MEASURE EVALUATION CRITERIA
<p>Structure – evidence that the measured structure supports the consistent delivery of effective processes or access that lead to improved health/avoidance of harm or cost/benefit.</p> <p>Patient experience – evidence that an association exists between the measure of patient experience of healthcare and the outcomes, values, and preferences of individuals/the public.</p> <p>Access – evidence that an association exists between access to a health service and the outcomes of, or experience with, care.</p> <p>Efficiency⁵ – demonstration of an association between the measured resource use and level of performance with respect to one or more of the other five IOM aims of quality.</p> <p><i>If not important to measure and report, STOP.</i></p>	
<p>2. Scientific acceptability of the measure properties: Extent to which the measure, as specified, produces consistent (reliable) and credible (valid) results about the quality of care when implemented.</p> <p>2a. The measure is well defined and precisely specified⁶ so that it can be implemented consistently within and across organizations and allow for comparability. The required data elements are of high quality as defined by NQF's Health Information Technology Expert Panel (HITEP).⁷</p>	<p>2. Scientific acceptability of the measure properties.</p> <p>2a. The composite measure is well defined and precisely specified so that it can be implemented consistently within and across organizations and allow for comparability. Composite specifications include methods for standardizing scales across component scores, scoring rules (i.e., how the component scores are combined or aggregated), weighting rules (i.e., whether all component scores are given equal or differential weighting when combined into the composite), handling of missing data, and required sample sizes.</p>

more

Table 1: Individual and Composite Measure Evaluation Criteria

The criteria for individual measure evaluation were updated with input from NQF Members, the public, and NQF's Consensus Standards Approval Committee and were approved by the NQF Board of Directors in August 2008.

INDIVIDUAL MEASURE EVALUATION CRITERIA	COMPOSITE MEASURE EVALUATION CRITERIA
<p>2b. Reliability testing⁸ demonstrates that the measure results are repeatable, producing the same results a high proportion of the time when assessed in the same population in the same time period.</p> <p>2c. Validity testing⁹ demonstrates that the measure reflects the quality of care provided, adequately distinguishing good and poor quality. If face validity is the only validity addressed, it is systematically assessed.</p> <p>2d. Clinically necessary measure exclusions are identified and must be: supported by evidence¹⁰ of sufficient frequency of occurrence so that results are distorted without the exclusion; <i>AND</i> a clinically appropriate exception (e.g., contraindication) to eligibility for the measure focus¹¹; <i>AND</i> precisely defined and specified: If there is substantial variability in exclusions across providers, the measure is specified so that exclusions are computable and the effect on the measure is transparent (i.e., impact clearly delineated, such as number of cases excluded, exclusion rates by type of exclusion). If patient preference (e.g., informed decisionmaking) is a basis for exclusion, there must be evidence that it strongly</p>	<p>2b. Reliability testing of the composite measure demonstrates that the results are repeatable, producing the same results a high proportion of the time when assessed in the same population in the same time period.</p> <p>2c. Validity testing demonstrates that the measure reflects the quality of care provided, adequately distinguishing good and poor quality. If face validity is the only validity addressed, it is systematically assessed.</p> <p>2f. Methods for scoring and analysis of the composite measure allow for identification of statistically significant and practically/clinically meaningful differences in performance.</p> <p>2h. If disparities in care have been identified, measure specifications, scoring, and analysis allow for identification of disparities through stratification of results (e.g., by race, ethnicity, socioeconomic status, gender); <i>OR</i> rationale/data justifies why stratification is not necessary or not feasible.</p> <p>New for composite. 2i. Component item/measure analysis (e.g., various correlation analyses such as internal consistency reliability), demonstrates that the included component items/measures fit the conceptual construct; <i>OR</i> justification and results for alternative analyses are provided.</p>

more

Table 1: Individual and Composite Measure Evaluation Criteria

The criteria for individual measure evaluation were updated with input from NQF Members, the public, and NQF's Consensus Standards Approval Committee and were approved by the NQF Board of Directors in August 2008.

INDIVIDUAL MEASURE EVALUATION CRITERIA	COMPOSITE MEASURE EVALUATION CRITERIA
<p>impacts performance on the measure, and the measure must be specified so that the information about patient preference and the effect on the measure is transparent¹² (e.g., numerator category computed separately, denominator exclusion category computed separately).</p> <p>2e. For outcome measures and other measures (e.g., resource use) when indicated: an evidence-based risk-adjustment strategy (e.g., risk models, risk stratification) is specified and is based on patient clinical factors that influence the measured outcome (but not disparities in care) and are present at start of care¹³</p> <p>OR</p> <p>rationale/data support no risk adjustment.</p> <p>2f. Data analysis demonstrates that methods for scoring and analysis of the specified measure allow for identification of statistically significant and practically/clinically meaningful¹⁴ differences in performance.</p> <p>2g. If multiple data sources/methods are allowed, there is demonstration that they produce comparable results.</p> <p>2h. If disparities in care have been identified, measure specifications, scoring, and analysis allow for identification of disparities through stratification of results (e.g., by race, ethnicity, socioeconomic status, gender);</p> <p>OR</p> <p>rationale/data justifies why stratification is not necessary or not feasible.</p>	<p>New for composite. 2j. Component item/measure analysis demonstrates that the included components contribute to the variation in the overall composite score;</p> <p>OR</p> <p>if not, justification for inclusion is provided.</p> <p>New for composite. 2k. The scoring/aggregation and weighting rules are consistent with the conceptual construct. (Simple, equal weighting is often preferred unless differential weighting is justified. Differential weights are determined by empirical analyses or a systematic assessment of expert opinion or values-based priorities.)</p> <p>New for composite. 2l. Analysis of missing component scores supports the specifications for scoring/aggregation and handling of missing component scores.</p>

more

Table 1: Individual and Composite Measure Evaluation Criteria

The criteria for individual measure evaluation were updated with input from NQF Members, the public, and NQF's Consensus Standards Approval Committee and were approved by the NQF Board of Directors in August 2008.

INDIVIDUAL MEASURE EVALUATION CRITERIA	COMPOSITE MEASURE EVALUATION CRITERIA
<p>3. Usability. Extent to which intended audiences (e.g., consumers, purchasers, providers, policymakers) can understand the results of the measure and are likely to find them useful for decisionmaking.</p> <p>3a. Demonstration that information produced by the measure is meaningful, understandable, and useful to the intended audience(s) for both public reporting (e.g., focus group, cognitive testing) and informing quality improvement (e.g., quality improvement initiatives).¹⁵ An important outcome that may not have an identified improvement strategy still can be useful for informing quality improvement by identifying the need for and stimulating new approaches to improvement.</p> <p>3b. The measure specifications are harmonized¹⁶ with other measures and are applicable to multiple levels and settings.</p> <p>3c. Review of existing endorsed measures and measure sets demonstrates that the measure provides a distinctive or additive value to existing NQF-endorsed measures (e.g., provides a more complete picture of quality for a particular condition or aspect of healthcare, is a more valid or efficient way to measure).</p>	<p>3. Usability</p> <p>3a. Demonstration that information produced by the composite measure is meaningful, understandable, and useful to the intended audience(s) for both public reporting (e.g., focus group, cognitive testing) and informing quality improvement (e.g., quality improvement initiatives).</p> <p>3b. The component measure specifications are harmonized.¹⁶</p> <p>3c. Review of existing endorsed measures and measure sets demonstrates that the composite measure provides a distinctive or additive value to existing NQF-endorsed measures (e.g., provides a more complete picture of quality for a particular condition or aspect of healthcare, is a more valid or efficient way to measure).</p> <p>New for composite. 3d. Data detail is maintained such that the composite measure can be decomposed into its components to facilitate transparency and understanding.</p> <p>New for composite. 3e. Demonstration (through pilot testing or operational data) that the composite measure achieves the stated purpose/objective.</p>

more

Table 1: Individual and Composite Measure Evaluation Criteria

The criteria for individual measure evaluation were updated with input from NQF Members, the public, and NQF's Consensus Standards Approval Committee and were approved by the NQF Board of Directors in August 2008.

INDIVIDUAL MEASURE EVALUATION CRITERIA	COMPOSITE MEASURE EVALUATION CRITERIA
<p>4. Feasibility. Extent to which the required data are readily available, retrievable without undue burden, and can be implemented for performance measurement.</p> <p>4a. For clinical measures, required data elements are routinely generated concurrent with and as a byproduct of care processes during care delivery.</p> <p>4b. The required data elements are available in electronic sources. If the required data are not in existing electronic sources, a credible, near-term path to electronic collection by most providers is specified, and clinical data elements are specified for transition to the electronic health record.</p> <p>4c. Exclusions should not require additional data sources beyond what is required for scoring the measure (e.g., numerator and denominator) unless justified as supporting measure validity.</p> <p>4d. Susceptibility to inaccuracies, errors, or unintended consequences and the ability to audit the data items to detect such problems are identified.</p> <p>4e. Demonstration that the data collection strategy (e.g., source, timing, frequency, sampling, patient confidentiality,¹⁷ etc.) can be implemented (e.g., already in operational use, or testing demonstrates that it is ready to put into operational use).</p>	<p>4. Feasibility</p> <p>4a. For clinical composite measures, overall the required data elements are routinely generated concurrent with and as a byproduct of care processes during care delivery.</p> <p>4b. The required data elements for the composite overall are available in electronic sources.</p> <p>4d. Susceptibility to inaccuracies, errors, or unintended consequences and the ability to audit the data items to detect such problems are identified.</p> <p>4e. Demonstration that the data collection strategy (e.g., source, timing, frequency, sampling, patient confidentiality, etc.) for obtaining all component measures can be implemented (e.g., already in operational use, or testing demonstrates that it is ready to put into operational use).</p>

more

Table 1: Individual and Composite Measure Evaluation Criteria

Notes

- 1 Examples of data on opportunity for improvement include but are not limited to prior studies, epidemiologic data, and measure data from pilot testing or implementation. If data are not available, the measure focus is systematically assessed (e.g., expert panel rating) and judged to be a quality problem.
- 2 Generally, rare event outcomes do not provide adequate information for improvement or discrimination; however, “never events” that are compared to zero are appropriate outcomes for public reporting and quality improvement.
- 3 The strength of the body of evidence for the specific measure focus should be systematically assessed and rated (e.g., the USPSTF grading system; see www.ahrq.gov/clinic/uspstf07/methods/benefit.htm). If the USPSTF grading system was not used, the grading system is explained, including how it relates to the USPSTF grades or why it does not. However, evidence is not limited to quantitative studies, and the best type of evidence depends upon the question being studied (e.g., randomized controlled trials appropriate for studying drug efficacy are not well suited for complex system changes). When qualitative studies are used, appropriate qualitative research criteria are used to judge the strength of the evidence.
- 4 Clinical care processes typically include multiple steps: assess → identify problem/potential problem → choose/plan intervention (with patient input) → provide intervention → evaluate impact on health status. If the measure focus is one step in such a multistep process, the step with the greatest effect on the desired outcome should be selected as the focus of measurement. For example, although assessment of immunization status and recommending immunization are necessary steps, they are not sufficient to achieve the desired impact on health status—patients must be vaccinated to achieve immunity. This does not preclude consideration of measures of preventive screening interventions where there is a strong link with desired outcomes (e.g., mammography) or measures for multiple care processes that affect a single outcome.
- 5 Efficiency of care is a measurement construct of cost of care or resource utilization associated with a specified level of quality of care. It is a measure of the relationship of the cost of care associated with a specific level of performance measured with respect to the other five IOM aims of quality. Efficiency might be thought of as a ratio, with quality as the numerator and cost as the denominator. As such, efficiency is directly proportional to quality and inversely proportional to cost. NQF’s *Measurement Framework: Evaluating Efficiency Across Episodes of Care* was posted for comment in November 2007 based on AQA Principles of Efficiency Measures at www.aqaalliance.org/files/PrinciplesofEfficiencyMeasurementApril2006.doc.
- 6 Measure specifications include the target population (e.g., denominator) to whom the measure applies, identification of those from the target population who achieved the specific measure focus (e.g., numerator), measurement time window, exclusions, risk adjustment, definitions, data elements, data source and instructions, sampling, and scoring/computation.
- 7 The HITEP criteria for high-quality data include: a) data are captured from an authoritative/accurate source; b) data are coded using recognized data standards; c) method of capturing data electronically fits the workflow of the authoritative source; d) data are available in EHRs; and e) data are auditable. NQF, *Health Information Technology Expert Panel Report: Recommended Common Data Types and Prioritized Performance Measures for Electronic Healthcare Information Systems*, Washington, DC: NQF; 2008.
- 8 Examples of reliability testing include but are not limited to inter-rater/abstractor or intrarater/abstractor studies; internal consistency for multi-item scales; and test-retest for survey items. Reliability testing may address the data items or final measure score.
- 9 Examples of validity testing include but are not limited to determining if measure scores adequately distinguish between providers known to have good or poor quality assessed by another valid method; correlation of measure scores with another valid indicator of quality for the specific topic; ability of measure scores to predict scores on some other related valid measure; and content validity for multi-item scales/tests. Face validity is a subjective assessment by experts of whether the measure reflects the quality of care (e.g., whether the proportion of patients with BP <140/90 is a marker of quality). If face validity is the only validity addressed, it is systematically assessed (e.g., ratings by relevant stakeholders), and the measure is judged to represent quality care for the specific topic and that the measure focus is the most important aspect of quality for the specific topic.
- 10 Examples of evidence that an exclusion distorts measure results include but are not limited to frequency of occurrence, sensitivity analyses with and without the exclusion, and variability of exclusions across providers.

more

Table 1: Individual and Composite Measure Evaluation Criteria Notes

-
- 11 Risk factors that influence outcomes should not be specified as exclusions.
 - 12 Patient preference is not a clinical exception to eligibility and can be influenced by provider interventions.
 - 13 Risk models should not obscure disparities in care for populations by including factors that are associated with differences/inequalities in care such as race, socioeconomic status, or gender (e.g., poorer treatment outcomes of African American men with prostate cancer, inequalities in treatment for CVD risk factors between men and women). It is preferable to stratify measures by race and socioeconomic status rather than to adjust out differences.
 - 14 With large enough sample sizes, small differences that are statistically significant may or may not be practically or clinically meaningful. The substantive question may be, for example, whether a statistically significant difference of one percentage point in the percentage of patients who received smoking cessation counseling (e.g., 74 percent versus 75 percent) is clinically meaningful, or whether a statistically significant difference of \$25 in cost for an episode of care (e.g., \$5,000 versus \$5,025) is practically meaningful. Measures with overall poor performance may not demonstrate much variability across providers.
 - 15 Public reporting and quality improvement are not limited to provider-level measures—community and population measures also are relevant for reporting and improvement.
 - 16 Measure harmonization refers to the standardization of specifications for similar measures on the same topic (e.g., *influenza immunization* of patients in hospitals or nursing homes), or related measures for the same target population (e.g., eye exam and HbA1c for *patients with diabetes*), or definitions applicable to many measures (e.g., age designation for children) so that they are uniform or compatible, unless differences are dictated by the evidence. The dimensions of harmonization can include numerator, denominator, exclusions, and data source and collection instructions. The extent of harmonization depends on the relationship of the measures, the evidence for the specific measure focus, and differences in data sources.
 - 17 All data collection must conform to laws regarding protected health information. Patient confidentiality is of particular concern with measures based on patient surveys and when there are small numbers of patients.
-

National Voluntary Consensus Standards for Mortality and Safety— Composite Measures

The development of the framework for evaluating composite measures was, in part, occasioned by the submission of four composite measures in response to an earlier NQF Call for Measures. For the reasons stated, this framework needed to be in place before moving forward with any future composite measure evaluation. It was determined that once the framework was developed and refined, it would be tested by using it to evaluate the four composite measures that were awaiting action. No further changes were made based on applying the evaluation framework to the measures. The measures are endorsed for public reporting and are expected to be useful for internal quality improvement.

Of note, reexamination of the component measures included in the composites was not carried out by the Steering Committee because the component measures had been evaluated by other NQF-constituted committees and were either NQF endorsed or recommended as components of the composite by the appropriate

Technical Advisory Panel under the NQF Hospital Care 2007 project. In the future, it is expected that composite measures will be evaluated by the committees most familiar with them—ideally, committees that evaluate the component measures.

All components of the four composite measures considered were selected by the developer, the Agency for Healthcare Research and Quality (AHRQ), from among its Quality Indicators, following a process that used an AHRQ expert panel to evaluate the measures before submitting them for evaluation by NQF.^Y Consistent with composite evaluation criterion 1.e., there was to be a clear and accepted rationale for selecting and approving components to ensure that composites proposed for endorsement meet accepted expectations. In the case of the composites advanced here, this information was provided and accepted. See Table 2 for the composite measures and subcomponents.

^Y The final reports of the AHRQ expert panels are available at:

www.qualityindicators.ahrq.gov/downloads/iqi/AHRQ_IQI_Workgroup_Final.pdf (mortality);

www.qualityindicators.ahrq.gov/downloads/pdi/AHRQ_PDI_Workgroup_Final.pdf (pediatric patient safety);

and www.qualityindicators.ahrq.gov/downloads/psi/AHRQ_PSI_Workgroup_Final.pdf (patient safety).

Table 2: National Voluntary Consensus Standards for Mortality and Safety: Composite Measures

COMPOSITE MEASURE TITLE	MEASURE ID ^a	COMPOSITE SUBCOMPONENTS AND AHRQ NUMBER	IP OWNER ^b
Mortality for selected conditions	0530	<ul style="list-style-type: none"> ■ Acute myocardial infarction (AMI) mortality (IQI 15) ■ Congestive heart failure (CHF) mortality (IQI 16) ■ Acute stroke mortality (IQI 17) ■ GI hemorrhage mortality (IQI 18) ■ Hip fracture mortality (IQI 19) ■ Pneumonia mortality (IQI 20) 	AHRQ
Pediatric patient safety for selected indicators	0532	<ul style="list-style-type: none"> ■ Accidental puncture or laceration (PDI 1) ■ Decubitus ulcer (PDI 2) ■ Iatrogenic pneumothorax (PDI 5) ■ Postoperative sepsis (PDI 10) ■ Postoperative wound dehiscence (PDI 11) ■ Selected infections due to medical care (PDI 12) 	AHRQ
Patient safety for selected indicators	0531	<ul style="list-style-type: none"> ■ Decubitus ulcer (PSI 3) ■ Iatrogenic pneumothorax (PSI 6) ■ Selected infections due to medical care (PSI 7) ■ Postoperative hip fracture (PSI 8) ■ Postoperative pulmonary embolism or deep vein thrombosis (PSI 12) ■ Postoperative sepsis (PSI 13) ■ Postoperative wound dehiscence (PSI 14) ■ Accidental puncture or laceration (PSI 15) 	AHRQ

^a Upon NQF endorsement, each measure receives a unique NQF measure ID number.

^b IP owner—intellectual property owner and copyright holder. For the most current specifications and supporting information, please refer to the IP owner, AHRQ - Agency for Healthcare Research and Quality (www.ahrq.gov).

Endorsed Composite Measures

NQF #0530 **Mortality for selected conditions** (AHRQ)

This composite measure includes all of the AHRQ Quality Indicators related to in-hospital mortality for specific conditions that are either NQF endorsed or assessed to be acceptable as components of the composite by the appropriate Technical Advisory Panel under the NQF Hospital Care Additional Priorities 2007 project.

Two key issues were considered in relation to this composite: 1) the utility of a composite when the individual condition mortality is known, which is most useful for improvement or healthcare decisions, and 2) covariate imbalance. For example, interhospital performance comparisons are problematic when such hospitals' composite mortality rates are based on markedly different distributions of conditions and patient severities, even when appropriate risk adjustment has been utilized.

Benefits of the mortality composites were proffered by the developer. The components are weighted based on the probability for each condition in order to maximize the outcome for the population. The measure would be of most benefit when the reason for the hospital admission was not known in advance. When a consumer does not know in advance for what hospitalization might be required or has competing interests, the weighted composite provides useful information to maximize population outcomes and, in this situation, to provide greater reliability. This rationale might apply, for example, when a health plan is

choosing hospitals to include in its network. When the reason for hospitalization is known in advance, the condition-specific rates would be most useful if that information is available. For example, AHRQ also reports on the component measures.

It was recommended to the developer that it clearly state the intended use and limitations of this composite. For example, it might be a useful measure for a health plan in choosing hospitals to include in its network (e.g., condition unknown, diagnosis-naïve). In this case, weighting based on the prevalence of conditions in the population would be appropriate and would maximize population benefit. The composite for mortality related to conditions should indicate that the intended purpose/objective is to apply this composite when patients' specific conditions are not known in advance of choosing a hospital.

NQF #0532 **Pediatric patient safety for selected indicators** (AHRQ)

The composite includes all of the AHRQ Quality Indicators related to in-hospital adverse events for the *pediatric* population that are either NQF endorsed or assessed to be acceptable as components of the composite by the appropriate Technical Advisory Panel under the NQF Hospital Care Additional Priorities 2007 project.

Key discussions that prefaced the endorsement recommendation for this measure related to the underlying concept of this composite. The developer proposed that it is appropriate to view the composite and its components from the standpoint of overall quality rather than

requiring tight causal linkages and the use of psychometric principles. With respect to the latter, it was recommended that justification should focus on the clear description of the purpose and quality construct and how the component measures fulfill the purpose and construct, as well as what is missing; why psychometric analysis was not deemed essential; an analysis of how they contribute to variability in the composite score; and how the scoring/aggregation achieves the purpose/quality construct.

Based on a determination that all components of the composite are of interest from a view of overall quality and the acceptability of the developer's response to questions, this composite measure was recommended for endorsement.

NQF #0531 Patient safety for selected indicators (AHRQ)

The composite includes all of the AHRQ Quality Indicators related to in-hospital adverse events for the *adult* population that are either NQF endorsed or assessed to be acceptable as components of the composite by the appropriate Technical Advisory Panel under the Hospital Care Additional Priorities 2007 project. The only NQF-endorsed indicator excluded from the composite is Death Among Surgical Inpatients with Serious Treatable Complications (NQF# 0200), because the indicator is materially different in structure and frequency of occurrence.

Composite Measure Not Endorsed

MORTALITY FOR SELECTED PROCEDURES (AHRQ) COM-002-08

The composite includes all of the AHRQ Quality Indicators related to the volume of specific procedures and in-hospital mortality for specific procedures that are either NQF endorsed or assessed to be acceptable as components of the composite by the appropriate Technical Advisory Panel under the NQF Hospital Care Additional Priorities 2007 project.

Two issues were discussed in relation to this composite: 1) the utility of a composite when the individual procedure mortality is known, which is most useful for improvement or healthcare decisions, and 2) covariate imbalance with different distributions of the procedures. The imbalance in the distribution of the procedures was viewed as a more significant issue for this measure than for the measure related to conditions. The determination to not recommend this measure for endorsement was based on the heterogeneity of cases across hospitals for the procedures, which was viewed as too great to comparatively represent quality of care in institutions as a composite measure.

Notes

- 1 Institute of Medicine, *Performance Measurement: Accelerating Improvement*, Washington, DC: National Academies Press; 2006.
- 2 Booyesen F, An overview and evaluation of composite indices of development, *Soc Indic Res*, 2002;59:115-151.
- 3 Fayers PM, Hand DJ, Causal variables, indicator variables and measurement scales: an example from quality of life, *J R Statist Soc A*, 2002;165(Part 2):233-261.
- 4 Kaplan SH, Normand SL, *Conceptual and Analytical Issues in Creating Composite Measures of Ambulatory Care Performance*, Washington, DC: National Quality Forum; 2006.
- 5 Nardo M, Saisana M, Saltelli A, Tarantola S, *Handbook on Constructing Composite Indicators: Methodology and User Guide*, OECD Statistics Working Paper, Paris: OECD Statistics Directorate; 2005. Report No.: STD/DOC(2005)3.
- 6 O'Brien SM, Shahian DM, DeLong ER, et al., Quality measurement in adult cardiac surgery: part 2—statistical considerations in composite measure scoring and provider rating, *Ann Thorac Surg*, 2007;83(Suppl 4):S13-S26.
- 7 Reeves D, Campbell SM, Adams J, et al., Combining multiple indicators of clinical quality: an evaluation of different analytic approaches, *Med Care*, 2007;45(6):489-496.
- 8 Shahian DM, Edwards FH, Ferraris VA, et al., Quality measurement in adult cardiac surgery: part 1—conceptual framework and measure selection, *Ann Thorac Surg*, 2007;83(Suppl 4):S3-S12.
- 9 Shahian DM, *Considerations in Composite Measure Development for Healthcare* (internal working document); 2007.
- 10 HealthPartners, *2006 Clinical Indicators Report: 2005/2006 Results*. Available at www.healthpartners.com/files/34432.pdf. Last accessed July 2008.
- 11 Nolan T, Berwick DM, All-or-none measurement raises the bar on performance, *JAMA*, 2006;295(10):1168-1170.
- 12 Jacobs R, Goddard M, Smith PC, How robust are hospital ranks based on composite performance measures? *Med Care*, 2005;43(12):1177-1184.
- 13 Nunnally JC, Bernstein IH, *Psychometric Theory*, 3rd ed., New York: McGraw-Hill; 1994.

Composite Measure Evaluation Framework and National Voluntary Consensus Standards for Mortality and Safety—Composite Measures: A Consensus Report

Appendix A

Specifications of the National Voluntary Consensus Standards for Mortality and Safety—Composite Measures

THE FOLLOWING TABLE PRESENTS the specifications for each of the NQF-endorsed[®] composite measures. This includes the specifications of each of the component measures. Component measures that are endorsed as individual measures are noted by an asterisk (*). All information presented has been derived directly from the measure developer without modification or alteration (except when the measure developer agreed to such modification during the NQF Consensus Development Process) and is current as of January 2009. The measures are open source, meaning they are fully accessible and disclosed.

The AHRQ composite measure workgroup reports for the measures (as well as for the measure not recommended) can be found online:

- Mortality for Selected Conditions—www.qualityindicators.ahrq.gov/downloads/iqi/AHRQ_IQI_Workgroup_Final.pdf
- Pediatric Patient Safety for Selected Indicators—www.qualityindicators.ahrq.gov/downloads/pdi/AHRQ_PDI_Workgroup_Final.pdf
- Patient Safety for Selected Indicators—www.qualityindicators.ahrq.gov/downloads/psi/AHRQ_PSI_Workgroup_Final.pdf

Appendix A – Specifications of the National Voluntary Consensus Standards for Mortality and Safety—Composite Measures

COMPOSITE MEASURE TITLE	IP OWNER ^a	NUMERATOR	DENOMINATOR	EXCLUSIONS	RISK ADJUSTMENT
Mortality for selected conditions NQF# 0530	AHRQ	Number of in-hospital deaths for acute myocardial infarction (AMI), congestive heart failure (CHF), acute stroke, gastrointestinal (GI) hemorrhage, hip fracture, and pneumonia (separately). See Inpatient Quality Indicators: Technical Specifications for additional details (available at www.qualityindicators.ahrq.gov/downloads/iqi/iqi_technical_specs_v32a.pdf).	Number of eligible adult discharges for AMI, CHF, acute stroke, GI hemorrhage, hip fracture, and pneumonia (separately). See Inpatient Quality Indicators: Technical Specifications for additional details (available at www.qualityindicators.ahrq.gov/downloads/iqi/iqi_technical_specs_v32a.pdf).	Component indicator specific.	Component indicator specific.
Mortality for Selected Conditions Component Measures					
Acute myocardial infarction mortality (AHRQ IQI 15) (recommended for individual endorsement)	AHRQ	Number of deaths (DISP=20) among cases meeting the inclusion and exclusion rules for the denominator.	Discharges, age 18 years and older, with an ICD-9-CM principal diagnosis code of AMI.	Cases: <ul style="list-style-type: none"> ■ missing discharge disposition (DISP=missing) ■ transferring to another short-term hospital (DISP=2). 	Age, sex, APR-DRG with risk-of-mortality subclass.
*Congestive heart failure mortality (AHRQ IQI 16, NQF# 0358)	AHRQ	Number of deaths (DISP=20) among cases meeting the inclusion and exclusion rules for the denominator.	Discharges, age 18 years and older, with an ICD-9-CM principal diagnosis code of CHF.	Cases: <ul style="list-style-type: none"> ■ missing discharge disposition (DISP=missing) ■ transferring to another short-term hospital (DISP=2) ■ in MDC 14 (pregnancy, childbirth, and puerperium) ■ in MDC 15 (newborns and other neonates). 	Age, sex, APR-DRG with risk-of-mortality subclass.

more

^a IP owner—intellectual property owner. ALL RIGHTS RESERVED. For the most current specifications and supporting information, please refer to the IP owner:
AHRQ - Agency for Healthcare Research and Quality (www.ahrq.gov)

Appendix A – Specifications of the National Voluntary Consensus Standards for Mortality and Safety—Composite Measures

COMPOSITE MEASURE TITLE	IP OWNER ^a	NUMERATOR	DENOMINATOR	EXCLUSIONS	RISK ADJUSTMENT
Mortality for Selected Conditions Component Measures					
*Acute stroke mortality (AHRQ IQI 17; NQF# 0467)	AHRQ	Number of deaths (DISP=20) among cases meeting the inclusion and exclusion rules for the denominator.	Discharges, age 18 years and older, with an ICD-9-CM principal diagnosis code of stroke.	<p>Cases:</p> <ul style="list-style-type: none"> ■ missing discharge disposition (DISP=missing) ■ transferring to another short-term hospital (DISP=2) ■ in MDC 14 (pregnancy, childbirth, and puerperium) ■ in MDC 15 (newborns and other neonates). 	Age, sex, APR-DRG with risk-of-mortality subclass.
GI hemorrhage mortality (AHRQ IQI 18) (for use in composite only)	AHRQ	Number of deaths (DISP=20) among cases meeting the inclusion and exclusion rules for the denominator.	Discharges, age 18 years and older, with an ICD-9-CM principal diagnosis code of gastrointestinal hemorrhage.	<p>Cases:</p> <ul style="list-style-type: none"> ■ missing discharge disposition (DISP=missing) ■ transferring to another short-term hospital (DISP=2) ■ in MDC 14 (pregnancy, childbirth, and puerperium) ■ in MDC 15 (newborns and other neonates). 	Age, sex, APR-DRG with risk-of-mortality subclass.
*Hip fracture mortality (AHRQ IQI 19; NQF# 0354)	AHRQ	Number of deaths (DISP=20) among cases meeting the inclusion and exclusion rules for the denominator.	Discharges, age 65 years and older, with an ICD-9-CM principal diagnosis code of hip fracture.	<p>Cases:</p> <ul style="list-style-type: none"> ■ with missing discharge disposition (DISP=missing) ■ transferring to another short-term hospital (DISP=2) ■ in MDC 14 (pregnancy, childbirth, and puerperium) ■ in MDC 15 (newborns and other neonates). 	Age, sex, APR-DRG with risk-of-mortality subclass.

more

Appendix A – Specifications of the National Voluntary Consensus Standards for Mortality and Safety—Composite Measures

COMPOSITE MEASURE TITLE	IP OWNER ^a	NUMERATOR	DENOMINATOR	EXCLUSIONS	RISK ADJUSTMENT
Mortality for Selected Conditions Component Measures					
<p>*Pneumonia mortality (AHRQ IQI 20; NQF# 0231)</p>	AHRQ	Number of deaths (DISP=20) among cases meeting the inclusion and exclusion rules for the denominator.	Discharges, age 18 years and older, with ICD-9-CM principal diagnosis code of pneumonia.	<p>Cases:</p> <ul style="list-style-type: none"> ■ with missing discharge disposition (DISP=missing) ■ transferring to another short-term hospital (DISP=2) ■ in MDC 14 (pregnancy, childbirth, and puerperium) ■ in MDC 15 (newborns and other neonates). 	Age, sex, APR-DRG with risk-of-mortality subclass.

more

Appendix A – Specifications of the National Voluntary Consensus Standards for Mortality and Safety—Composite Measures

COMPOSITE MEASURE TITLE	IP OWNER ^a	NUMERATOR	DENOMINATOR	EXCLUSIONS	RISK ADJUSTMENT
Pediatric Patient Safety for Selected Indicators Component Measures					
<i>Note: Appendices noted in the specifications below may be accessed in the AHRQ PDI Technical Specifications Version 3.2 (March 10, 2008), available at www.qualityindicators.ahrq.gov/downloads/pdi/pdi_technical%20specs_v31.pdf.</i>					
Pediatric patient safety for selected indicators NQF# 0532	AHRQ	Number of potentially preventable adverse events for accidental puncture or laceration, decubitus ulcer, iatrogenic pneumothorax, postoperative wound dehiscence, postoperative sepsis, and selected infection due to medical care (separately). See Pediatric Quality Indicators: Technical Specifications for additional details, including appendices (available at www.qualityindicators.ahrq.gov/downloads/pdi/pdi_technical_specs_v32.pdf).	Number of eligible pediatric discharges for accidental puncture or laceration, decubitus ulcer, iatrogenic pneumothorax, postoperative wound dehiscence, postoperative sepsis, and selected infection due to medical care (separately). See Pediatric Quality Indicators: Technical Specifications for additional details, including appendices (available at www.qualityindicators.ahrq.gov/downloads/pdi/pdi_technical_specs_v32.pdf).	Component indicator specific.	Component indicator specific.
*Accidental puncture or laceration (AHRQ PDI 1; NQF# 0344)	AHRQ	Discharges among cases meeting the inclusion and exclusion rules for the denominator with ICD-9-CM codes denoting accidental cut, puncture, perforation, or laceration during a procedure in any secondary diagnosis field.	All surgical and medical discharges under age 18 defined by specific Surgical and Medical Diagnosis Related Group (DRG). See Appendix B: Surgical Discharge DRGs. See Appendix E: Medical Discharge DRGs.	Cases: <ul style="list-style-type: none"> ■ with ICD-9-CM code denoting accidental cut, puncture, perforation, or laceration in the principal diagnosis field (or secondary diagnosis code if present on admission) ■ in MDC 14 (pregnancy, childbirth, and puerperium) ■ with ICD-9-CM procedure code for spine surgery ■ normal newborn (DRG 391) (According to developer, exclude normal newborns because they do not usually undergo procedures that put normal newborns at risk for these complications) ■ newborns with birthweight less than 500 grams. See Appendix G: Low Birth Weight Categories.	Age in days, sex, neonate weight, DRG, comorbidity categories, procedure class.

more

Appendix A – Specifications of the National Voluntary Consensus Standards for Mortality and Safety—Composite Measures

COMPOSITE MEASURE TITLE	IP OWNER ^a	NUMERATOR	DENOMINATOR	EXCLUSIONS	RISK ADJUSTMENT
Pediatric Patient Safety for Selected Indicators Component Measures					
<i>Note: Appendices noted in the specifications below may be accessed in the AHRQ PDI Technical Specifications Version 3.2 (March 10, 2008), available at www.qualityindicators.ahrq.gov/downloads/pdi/pdi_technical%20specs_v31.pdf.</i>					
<p>*Decubitus ulcer (AHRQ PDI 2; NQF# 0377)</p>	AHRQ	Discharges among cases meeting the inclusion and exclusion rules for the denominator with ICD-9-CM codes of decubitus ulcer in any secondary diagnosis field.	<p>All surgical and medical discharges under age 18 defined by specific Surgical and Medical DRG.</p> <p>See Appendix B: Surgical Discharge DRGs. See Appendix E: Medical Discharge DRGs.</p>	<p>Cases:</p> <ul style="list-style-type: none"> ■ with ICD-9-CM code of decubitus ulcer in the principal diagnosis field (or secondary diagnosis field if present on admission) ■ in MDC 9 (skin, subcutaneous tissue, and breast) ■ in MDC 14 (pregnancy, childbirth, and puerperium) ■ admitted from a long-term care facility (SID Admission Source=3) or transferred from an acute care facility (SID Admission Source=2) ■ neonates ■ with length of stay of less than 5 days. <p>See Appendix A: Operating Room Procedure Codes.</p>	Age in days, sex, neonate weight, DRG, comorbidity categories, risk class.

more

Appendix A – Specifications of the National Voluntary Consensus Standards for Mortality and Safety—Composite Measures

COMPOSITE MEASURE TITLE	IP OWNER ^a	NUMERATOR	DENOMINATOR	EXCLUSIONS	RISK ADJUSTMENT
Pediatric Patient Safety for Selected Indicators Component Measures					
<i>Note: Appendices noted in the specifications below may be accessed in the AHRQ PDI Technical Specifications Version 3.2 (March 10, 2008), available at www.qualityindicators.ahrq.gov/downloads/pdi/pdi_technical%20specs_v31.pdf.</i>					
<p>*Iatrogenic pneumothorax (AHRQ PDI 5; NQF# 0348)</p>	AHRQ	Discharges among cases meeting the inclusion and exclusion rules for the denominator with ICD-9-CM code of 512.1 in any secondary diagnosis field.	<p>All surgical and medical discharges under age 18 defined by specific Surgical and Medical DRG.</p> <p>See Appendix B: Surgical Discharge DRGs. See Appendix E: Medical Discharge DRGs.</p>	<p>Cases:</p> <ul style="list-style-type: none"> ■ with ICD-9-CM code of 512.1 in the principal diagnosis field (or secondary diagnosis field present on admission) ■ in MDC 14 (pregnancy, childbirth, and puerperium) ■ with ICD-9-CM diagnosis code of chest trauma or pleural effusion ■ with ICD-9-CM procedure code of thoracic surgery, lung or pleural biopsy or diaphragmatic surgery repair or assigned to a cardiac surgery DRG ■ normal newborn (DRG 391) ■ neonates with birthweight less than 2500 grams <p>See Appendix G: Low Birth Weight Categories.</p>	Age in days, sex, neonate weight, DRG, comorbidity categories.

more

Appendix A – Specifications of the National Voluntary Consensus Standards for Mortality and Safety—Composite Measures

COMPOSITE MEASURE TITLE	IP OWNER ^a	NUMERATOR	DENOMINATOR	EXCLUSIONS	RISK ADJUSTMENT
Pediatric Patient Safety for Selected Indicators Component Measures					
<i>Note: Appendices noted in the specifications below may be accessed in the AHRQ PDI Technical Specifications Version 3.2 (March 10, 2008), available at www.qualityindicators.ahrq.gov/downloads/pdi/pdi_technical%20specs_v31.pdf.</i>					
<p>Postoperative sepsis (PDI 10) (for use in composite only)</p>	AHRQ	<p>Discharges among cases meeting the inclusion and exclusion rules for the denominator with ICD-9-CM code for sepsis in any secondary diagnosis field.</p> <p>See Appendix F: Sepsis Diagnosis Codes.</p>	<p>All surgical discharges under age 18 defined by specific Surgical DRGs and an ICD-9-CM code for an operating room procedure.</p> <p>See Appendix A: Operating Room Procedure Codes.</p> <p>See Appendix B: Surgical Discharge DRGs.</p>	<p>Cases:</p> <ul style="list-style-type: none"> ■ with ICD-9-CM codes for sepsis in the principal diagnosis field (or secondary diagnosis field if present on admission) ■ in MDC 14 (pregnancy, childbirth, and puerperium) ■ with ICD-9-CM codes for infection in the principal diagnosis field (or secondary diagnosis field if present on admission) ■ neonates ■ with length of stay of less than 4 days ■ with DRG code in surgical class 4 <p>See Appendix F: Sepsis Diagnosis Codes. See Appendix H: Infection Diagnosis Codes.</p>	<p>Age in days, sex, neonate weight, DRG, comorbidity categories, procedure type, risk class.</p>

more

Appendix A – Specifications of the National Voluntary Consensus Standards for Mortality and Safety—Composite Measures

COMPOSITE MEASURE TITLE	IP OWNER ^a	NUMERATOR	DENOMINATOR	EXCLUSIONS	RISK ADJUSTMENT
Pediatric Patient Safety for Selected Indicators Component Measures					
<i>Note: Appendices noted in the specifications below may be accessed in the AHRQ PDI Technical Specifications Version 3.2 (March 10, 2008), available at www.qualityindicators.ahrq.gov/downloads/pdi/pdi_technical%20specs_v31.pdf.</i>					
<p>*Postoperative wound dehiscence (AHRQ PDI 11; NQF# 0367)</p>	AHRQ	Discharges among cases meeting the inclusion and exclusion rules for the denominator with ICD-9-CM procedure code for reclosure of postoperative disruption of abdominal wall (54.61) in any procedure field.	All abdominopelvic surgical discharges under age 18 defined by ICD-9-CM code for an abdominopelvic procedure.	<p>Cases:</p> <ul style="list-style-type: none"> ■ where ICD-9-CM procedure code for reclosure of postoperative disruption of abdominal wall occurs before or on the same day as the first abdominopelvic surgery procedure ■ in MDC 14 (pregnancy, childbirth, and puerperium) ■ with ICD-9-CM code for high- or intermediate-risk immunocompromised state in any diagnosis field ■ with ICD-9-CM code for gastroschisis or umbilical hernia repair in newborns (omphalacele repair) performed before reclosure ■ neonates with birthweight less than 500 grams ■ with length of stay less than 2 days. <p>See Appendix C: ICD-9-CM codes for High-risk Immunocompromised States. See Appendix D: ICD-9-CM Codes for Intermediate-risk Immunocompromised States. See Appendix G: Low Birth Weight Categories.</p>	Age in days, sex, neonate weight, DRG, comorbidity categories, procedure type.

more

Appendix A – Specifications of the National Voluntary Consensus Standards for Mortality and Safety—Composite Measures

COMPOSITE MEASURE TITLE	IP OWNER ^a	NUMERATOR	DENOMINATOR	EXCLUSIONS	RISK ADJUSTMENT
Pediatric Patient Safety for Selected Indicators Component Measures					
<i>Note: Appendices noted in the specifications below may be accessed in the AHRQ PDI Technical Specifications Version 3.2 (March 10, 2008), available at www.qualityindicators.ahrq.gov/downloads/pdi/pdi_technical%20specs_v31.pdf.</i>					
Selected infections due to medical care (AHRQ PDI 12) (for use in composite only)	AHRQ	Discharges among cases meeting the inclusion and exclusion rules for the denominator with ICD-9-CM code of 999.3, 999.31, or 996.62 in any secondary diagnosis field.	All surgical and medical discharges under age 18 defined by specific surgical and medical DRG. See Appendix B: Surgical Discharge DRGs. See Appendix E: Medical Discharge DRGs.	Cases: <ul style="list-style-type: none"> ■ with ICD-9-CM code of 999.3, 999.31, or 996.62 in the principal diagnosis field (or secondary diagnosis field if present on admission) ■ in MDC 14 (pregnancy, childbirth, and puerperium) ■ newborns ■ neonates with birthweight less than 500 grams ■ with length of stay less than 2 days. 	Age in days, sex, neonate weight, DRG, comorbidity categories, risk class.
Patient safety for selected indicators NQF# 0531	AHRQ	Number of potentially preventable adverse events for decubitus ulcer, iatrogenic pneumothorax, selected infections due to medical care, postoperative hip fracture, postoperative deep vein thrombosis (DVT) or pulmonary embolism (PE), postoperative sepsis, postoperative wound dehiscence, and accidental puncture or laceration (separately). See Patient Safety Indicators: Technical Specifications for additional details, including appendices (available at www.qualityindicators.ahrq.gov/downloads/psi/psi_technical_specs_v32.pdf).	Number of eligible adult discharges for decubitus ulcer, iatrogenic pneumothorax, selected infections due to medical care, postoperative hip fracture, postoperative DVT or PE, postoperative sepsis, postoperative wound dehiscence, and accidental puncture or laceration (separately). See Patient Safety Indicators: Technical Specifications for additional details, including appendices (available at www.qualityindicators.ahrq.gov/downloads/psi/psi_technical_specs_v32.pdf).	Component indicator specific.	Component indicator specific.

more

Appendix A – Specifications of the National Voluntary Consensus Standards for Mortality and Safety—Composite Measures

COMPOSITE MEASURE TITLE	IP OWNER ^a	NUMERATOR	DENOMINATOR	EXCLUSIONS	RISK ADJUSTMENT
Patient Safety for Selected Indicators Component Measures					
<p>Decubitus ulcer (AHRQ PSI 3) (for use in composite only)</p>	AHRQ	Discharges with ICD-9-CM code of decubitus ulcer in any secondary diagnosis field among cases meeting the inclusion and exclusion rules for the denominator.	All medical and surgical discharges age 18 years and older defined by specific DRG. See Appendix B: Surgical Discharge DRGs. See Appendix F: Medical Discharge DRGs.	<p>Cases:</p> <ul style="list-style-type: none"> ■ with ICD-9-CM code of decubitus ulcer in the principal diagnosis field (or in a secondary diagnosis field if present on admission) ■ in MDC 9 (skin, subcutaneous tissue, and breast) ■ in MDC 14 (pregnancy, childbirth, and puerperium) ■ with ICD-9-CM diagnosis code of hemiplegia, paraplegia, or quadriplegia ■ with ICD-9-CM diagnosis code of spina bifida or anoxic brain damage ■ with ICD-9-CM procedure code for debridement or pedicle graft before or on the same day as the major operating room procedure (surgical cases only) ■ admitted from a long-term care facility (SID Admission Source=3) or transferred from an acute care facility (SID Admission Source=2) ■ with length of stay of less than 5 days. <p>See Appendix A: Operating Room Procedure.</p>	Age, sex, DRG, comorbidity categories.

more

Appendix A – Specifications of the National Voluntary Consensus Standards for Mortality and Safety—Composite Measures

COMPOSITE MEASURE TITLE	IP OWNER ^a	NUMERATOR	DENOMINATOR	EXCLUSIONS	RISK ADJUSTMENT
Patient Safety for Selected Indicators Component Measures					
<p>*Iatrogenic pneumothorax (AHRQ PSI 6; NQF# 0346)</p>	AHRQ	Discharges with ICD-9-CM code of 512.1 in any secondary diagnosis field among cases meeting the inclusion and exclusion rules for the denominator.	All surgical and medical discharges age 18 and older defined by specific DRG. See Appendix B: Surgical Discharge DRGs. See Appendix F: Medical Discharge DRGs.	<p>Cases:</p> <ul style="list-style-type: none"> ■ with ICD-9-CM code of 512.1 in the principal diagnosis field (or secondary diagnosis field if present on admission) ■ in MDC 14 (pregnancy, childbirth, and puerperium) ■ with ICD-9-CM diagnosis code of chest trauma or pleural effusion ■ with ICD-9-CM procedure code for diaphragmatic surgery repair ■ with ICD-9-CM procedure code for thoracic surgery, lung or pleural biopsy, or assigned to cardiac surgery DRGs. 	Age, sex, DRG, comorbidity categories.
<p>Selected infections due to medical care (AHRQ PSI 7) (for use in composite only)</p>	AHRQ	Discharges with ICD-9-CM code of 999.3, 999.31, or 996.62 in any secondary diagnosis field among cases meeting the inclusion and exclusion rules for the denominator.	All surgical and medical discharges age 18 and older defined by specific DRG. See Appendix B: Surgical Discharge DRGs. See Appendix F: Medical Discharge DRGs.	<p>Cases:</p> <ul style="list-style-type: none"> ■ with ICD-9-CM code of 999.3, 999.31, or 996.62 in the principal diagnosis field (or secondary diagnosis field if present on admission) ■ with ICD-9-CM diagnosis code for immunocompromised state or cancer or assigned to cancer DRGs ■ with length of stay less than 2 days. <p>See Appendix D: ICD-9-CM Codes for Immunocompromised States. See Appendix E: Cancer Codes. See Appendix Q: Cancer DRGs.</p>	Age, sex, DRG, comorbidity categories.

more

Appendix A – Specifications of the National Voluntary Consensus Standards for Mortality and Safety—Composite Measures

COMPOSITE MEASURE TITLE	IP OWNER ^a	NUMERATOR	DENOMINATOR	EXCLUSIONS	RISK ADJUSTMENT
Patient Safety for Selected Indicators Component Measures					
<p>Postoperative hip fracture (AHRQ PSI 8) (for use in composite only)</p>	AHRQ	Discharges with ICD-9-CM code for hip fracture in any secondary diagnosis field among cases meeting the inclusion and exclusion rules for the denominator.	<p>All surgical discharges age 18 years and older defined by specific DRG and an ICD-9-CM code for an operating room procedure.</p> <p>See Appendix A: Operating Room Procedure Codes.</p> <p>See Appendix B: Surgical Discharge DRGs.</p>	<p>Cases:</p> <ul style="list-style-type: none"> ■ with ICD-9-CM code of hip fracture in the principal diagnosis field (or secondary diagnosis field if present on admission) ■ in MDC 8 (diseases and disorders of the musculoskeletal system and connective tissue) ■ in MDC 14 (pregnancy, childbirth, and puerperium) ■ where the only operating room procedure is hip fracture repair ■ where a procedure for hip fracture repair occurs before or on the same day as the first operating room procedure ■ ICD-9-CM code of seizure, syncope, stroke, coma, cardiac arrest, poisoning, trauma, delirium and other psychoses, or anoxic brain injury in the principal diagnosis field (or secondary diagnosis field if present on admission) ■ ICD-9-CM diagnosis of metastatic cancer, lymphoid malignancy or bone malignancy, or self-inflicted injury. <p>See Appendix C: ICD-9-CM Trauma Diagnosis Codes.</p> <p>See Appendix J: Trauma DRGs.</p> <p>See Appendix L: Self-Inflicted Injury Diagnosis Codes.</p>	Age, sex, DRG, comorbidity categories.

more

Appendix A – Specifications of the National Voluntary Consensus Standards for Mortality and Safety—Composite Measures

COMPOSITE MEASURE TITLE	IP OWNER ^a	NUMERATOR	DENOMINATOR	EXCLUSIONS	RISK ADJUSTMENT
Patient Safety for Selected Indicators Component Measures					
<p>Postoperative pulmonary embolism or deep vein thrombosis (AHRQ PSI 12; NQF# 0450)</p>	AHRQ	<p>Discharges among cases meeting the inclusion and exclusion rules for the denominator with ICD-9-CM codes for deep vein thrombosis or pulmonary embolism in any secondary diagnosis field.</p> <p>See Appendix K: Pulmonary Embolism/Deep Vein Thrombosis Diagnosis Codes.</p>	<p>All surgical discharges age 18 and older defined by specific DRG and ICD-9-CM code for an operating room procedure.</p> <p>See Appendix A: Operating Room Procedure Codes.</p> <p>See Appendix B: Surgical Discharge DRGs.</p>	<p>Cases:</p> <ul style="list-style-type: none"> ■ with ICD-9-CM code for deep vein thrombosis or pulmonary embolism in the principal diagnosis field (or secondary diagnosis field if present on admission) ■ in MDC 14 (pregnancy, childbirth, and puerperium) ■ where a procedure for interruption of vena cava is the only operating room procedure, where a procedure for interruption of vena cava occurs before or on the same day as the first operating room procedure. 	Age, sex, DRG, comorbidity categories.

more

Appendix A – Specifications of the National Voluntary Consensus Standards for Mortality and Safety—Composite Measures

COMPOSITE MEASURE TITLE	IP OWNER ^a	NUMERATOR	DENOMINATOR	EXCLUSIONS	RISK ADJUSTMENT
Patient Safety for Selected Indicators Component Measures					
<p>Postoperative sepsis (AHRQ PSI 13) (for use in composite only)</p>	AHRQ	<p>Discharges among cases meeting the inclusion and exclusion rules for the denominator with ICD-9-CM code for sepsis in any secondary diagnosis field.</p> <p>See Appendix O: Sepsis Diagnosis Codes.</p>	<p>All elective* surgical discharges age 18 and older defined by specific DRG and ICD-9-CM code for an operating room procedure.</p> <p>See Appendix A: Operating Room Procedure Codes.</p> <p>See Appendix B: Surgical Discharge DRGs.</p> <p><i>*Elective—SID Admission Type=3.</i></p>	<p>Cases:</p> <ul style="list-style-type: none"> ■ with ICD-9-CM code of sepsis in the principal diagnosis field (or secondary diagnosis field if present on admission) ■ in MDC 14 (pregnancy, childbirth, and puerperium) ■ with ICD-9-CM code of infection in the principal diagnosis field (or secondary diagnosis field if present on admission) ■ with ICD-9-CM diagnosis code for immunocompromised state or cancer ■ with a length of stay of less than 4 days. <p>See Appendix P: Infection Diagnosis Codes and DRGs.</p> <p>See Appendix D: ICD-9-CM Codes for Immunocompromised States.</p> <p>See Appendix E: Cancer Codes.</p> <p>See Appendix Q: Cancer DRGs.</p>	Age, sex, DRG, comorbidity categories.

more

Appendix A – Specifications of the National Voluntary Consensus Standards for Mortality and Safety—Composite Measures

COMPOSITE MEASURE TITLE	IP OWNER ^a	NUMERATOR	DENOMINATOR	EXCLUSIONS	RISK ADJUSTMENT
Patient Safety for Selected Indicators Component Measures					
<p>*Postoperative wound dehiscence (AHRQ PSI 14; NQF# 0368)</p>	AHRQ	Discharges among cases meeting the inclusion and exclusion rules for the denominator with ICD-9-CM code for reclosure of postoperative disruption of abdominal wall (54.61) in any procedure field.	All abdominopelvic surgical discharges age 18 and older defined by ICD-9-CM procedure code.	<p>Cases:</p> <ul style="list-style-type: none"> ■ where a procedure for reclosure of postoperative disruption of abdominal wall occurs before or on the same day as the first abdominopelvic surgery procedure ■ in MDC 14 (pregnancy, childbirth, and puerperium) ■ with ICD-9-CM diagnosis code of immunocompromised state ■ where length of stay is less than 2 days. <p>See Appendix D: ICD-9-CM Codes for Immunocompromised States.</p>	Age, sex, DRG, comorbidity categories.
<p>*Accidental puncture or laceration (AHRQ PSI 15; NQF# 0345)</p>	AHRQ	Discharges among cases meeting the inclusion and exclusion rules for the denominator with ICD-9-CM code denoting accidental cut, puncture, perforation, or laceration during a procedure in any secondary diagnosis field.	All surgical and medical discharges age 18 years and older defined by specific DRG. See Appendix B: Surgical Discharge DRGs. See Appendix F: Medical Discharge DRGs.	<p>Cases:</p> <ul style="list-style-type: none"> ■ with ICD-9-CM code denoting accidental cut, puncture, perforation, or laceration during a procedure in the principal diagnosis field (or secondary diagnosis field if present on admission) ■ in MDC 14 (pregnancy, childbirth, and puerperium) ■ with ICD-9-CM procedure code for spine surgery. 	Age, sex, DRG, comorbidity categories.

Composite Measure Evaluation Framework and National Voluntary Consensus Standards for Mortality and Safety—Composite Measures: A Consensus Report

Appendix B Glossary

Term	Definition
All-or-none scoring	A percentage is determined by applying an all-or-none rule at the patient level. The denominator could be the number of patients eligible to receive at least one of the identified elements of care, and the numerator could be the number of patients who actually received all of the care for which the specific patient was eligible. No partial credit is given. ¹
Bundle	A series of interventions related to a specific condition that, when implemented together, will achieve significantly better outcomes than when implemented individually. This term was developed by faculty at the Institute for Healthcare Improvement. See www.ihl.org/IHI/Topics/CriticalCare/IntensiveCare/ImprovementStories/BundleUpforSafety.htm .
Clinimetric approach	Approach to developing a scale that relies on the required relationships between the observed items and the attribute for which an index is being defined. The most important attributes to be included in the index are not expected to be homogeneous because they indicate different aspects of a complex clinical phenomenon. ²
Component	A constituent part or element of a composite measure.
Composite measure	A combination of two or more individual measures into a single measure that results in a single score.
Construct	An abstract phenomenon that is measured indirectly through less abstract indicators.
Domain	A dimension or aspect of a construct.

Indicator	Sometimes used interchangeably with measure, but may indicate a more descriptive level than the term “measure,” which indicates the operational definition.
Item	A single question on a measurement scale or instrument.
Latent variable	An unobserved trait or characteristic.
Measure	Numeric quantification of some concept. A quality measure is a numeric quantification of healthcare quality.
Opportunity scoring	Scoring used with process measures, determined from the sum of all numerators (achieved the desired process) divided by the sum of all denominators (i.e., number of eligible patients or opportunities, which could vary by measure).
Paired measures	Individual measures that should be measured concurrently in the same population; however, the results are not combined into a single score.
Psychometric approach	Approach to developing a scale that relies on the relationships between the items that have been measured where the multiple component items are all measuring more or less the same single attribute.
Scale	A measure of an attribute composed of a set of related items. A score on the scale represents a point along a continuum representing more or less of the attribute.
Subscale	A measure of a dimension of a scale composed of a subset of the items in a scale.
Variable	A characteristic or attribute that varies within and among people or the subjects of study.

Notes

- 1 Nolan T, Berwick DM, All-or-none measurement raises the bar on performance, *JAMA*, 2006;295(10):1168-1170.
- 2 Fayers PM, Hand DJ, Causal variables, indicator variables and measurement scales: an example from quality of life, *J R Statist Soc A*, 2002;165(Part 2):233-261.

THE NATIONAL QUALITY FORUM (NQF) is a private, nonprofit, open membership, public benefit corporation whose mission is to improve the American healthcare system so that it can be counted on to provide safe, timely, compassionate, and accountable care using the best current knowledge. Established in 1999, NQF is a unique public-private partnership having broad participation from all parts of the healthcare industry. As a voluntary consensus standard-setting organization, NQF seeks to develop a common vision for healthcare quality improvement, create a foundation for standardized healthcare performance data collection and reporting, and identify a national strategy for healthcare quality improvement. NQF provides an equitable mechanism for addressing the disparate priorities of healthcare's many stakeholders.

NATIONAL QUALITY FORUM
601 13th Street NW
Suite 500 North
Washington, DC 20005
202-783-1300
www.qualityforum.org

